

This Page Is Inserted by IFW Operations  
and is not a part of the Official Record

## **BEST AVAILABLE IMAGES**

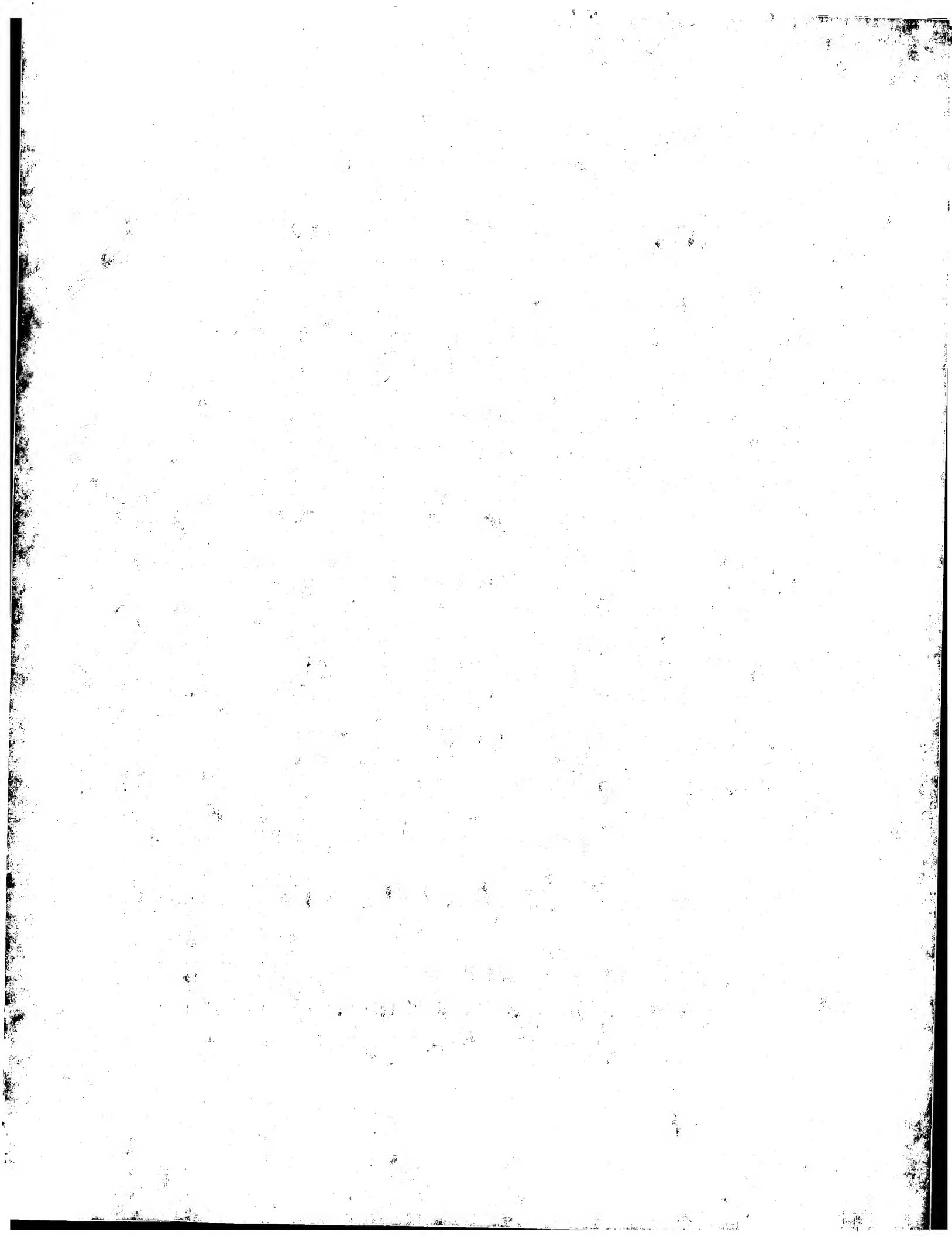
Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning documents *will not* correct images,  
please do not report the images to the  
Image Problem Mailbox.**

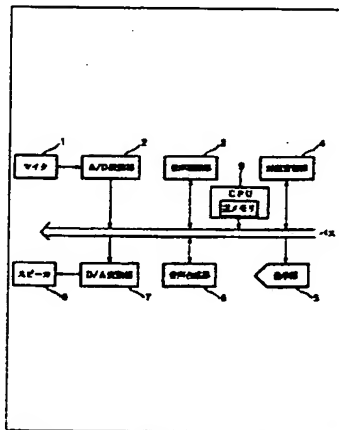


WPI

- TI - Information processing device for audio interaction system - in which output control part suspends concerned output during interruption and decides whether to respond to new input or to continue with suspended output
- AB - J08146991 The device inputs an audio signal through a microphone (1). The input audio signal is converted to a digital signal in an A/D conversion part (2). The digital signal is recognised as an audio signal in an audio recognition part (3). An audio synthesis part (6) outputs an audio signal in accordance with the audio signal recognised by the audio signal recognition part.
- The audio signal output is given to a speaker (8) and displayed on a display part (5). An output control part controls the output from the speaker. When a user utters during this time, the concerned output is suspended temporarily. Then, the output control part decides whether to respond to the new input or to continue with the suspended output.
  - ADVANTAGE - Provides required information quickly.
  - (Dwg.2/8)
- PN - JP8146991 A 19960607 DW199633 G10L3/00 008pp
- PR - JP19940283259 19941117
- PA - (CANO ) CANON KK
- MC - T01-C08A W04-V04 W04-V04A1
- DC - P86 T01 W04
- IC - G10L3/00
- AN - 1996-325401 [33]

PAJ

- TI - INFORMATION PROCESSOR AND ITS CONTROL METHOD
- AB - PURPOSE: To make a conversion by speech input natural and speedily obtain information which is necessary for a user.
- CONSTITUTION: A speech inputted from a microphone 1 is converted into a digital signal through an A/D conversion part 2 and recognized by a speech recognition part 3. An answer sentence based upon the recognition result is generated and made into a speech signal by a speech synthesis part 6, and the speech signal is displayed at a display part 5. If the user speaks during the said output, the output is interrupted temporarily and whether the output is restarted or ended is determined according to the speech contents.
- PN - JP8146991 A 19960607
- PD - 1996-06-07
- ABD - 19961031
- ABV - 199610
- AP - JP19940283259 19941117
- PA - CANON INC
- IN - KAWASAKI KATSUHIKO;KOMORI YASUHIRO;OHORA YASUNORI
- I - G10L3/00



<First Page Image>



(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平8-146991

(43) 公開日 平成8年(1996)6月7日

(51) Int.Cl.<sup>6</sup>

G 1 0 L 3/00

識別記号

5 7 1 J

庁内整理番号

H

F I

技術表示箇所

審査請求 未請求 請求項の数12 O L (全 8 頁)

(21) 出願番号

特願平6-283259

(22) 出願日

平成6年(1994)11月17日

(71) 出願人 000001007

キヤノン株式会社

東京都大田区下丸子3丁目30番2号

(72) 発明者 川崎 勝彦

東京都大田区下丸子3丁目30番2号 キヤ

ノン株式会社内

(72) 発明者 小森 康弘

東京都大田区下丸子3丁目30番2号 キヤ

ノン株式会社内

(72) 発明者 大洞 恭則

東京都大田区下丸子3丁目30番2号 キヤ

ノン株式会社内

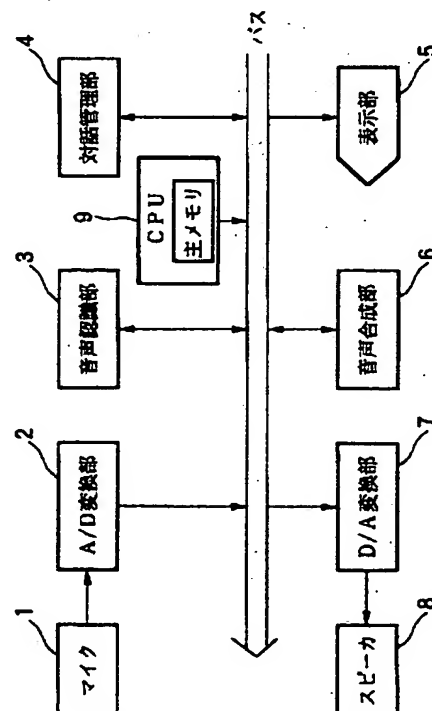
(74) 代理人 弁理士 大塚 康徳 (外1名)

(54) 【発明の名称】 情報処理装置及びその制御方法

(57) 【要約】

【目的】 音声入力による対話を自然な状態にさせると共に、尚且つ、ユーザにとっては必要な情報を迅速に得ることを可能にする。

【構成】 マイク1から入力された音声をA/D変換部2を介してデジタル信号にした後、音声認識部3で認識する。この認識した結果に基づく応答文を作成し、それを音声合成部6で音声信号にすると共に、それを表示部5に表示する。この出力中に、ユーザが発声すると、その出力を一時的に中断し、その発声内容に応じて再開するか、終了するかを決定する。



1

## 【特許請求の範囲】

【請求項1】 音声入力手段と、音声認識手段と、前記音声入力手段に対して応答する出力手段を備えた情報処理装置において、

前記音声入力手段から入力された音声を認識し、音声入力に対応する文を生成する文生成手段と、  
該文生成手段で生成された文に対応する出力文を構成し、前記出力手段を制御して外部に出力する出力制御手段と、

該出力制御手段で出力中、前記音声入力手段からの入力を監視し、音声入力があった場合に、当該出力を一時的に停止する第1の制御手段と、

入力された音声に応じて、出力途中であった応答文を再開するか、中断するかを制御する第2の制御手段とを備えることを特徴とする情報処理装置。

【請求項2】 前記出力手段は、音声出力手段及び対話文を表示する表示手段の両方を含むことを特徴とする請求項第1項に記載の情報処理装置。

【請求項3】 前記第2の制御手段は、音声入力手段から入力された音声情報の長さによって制御することを特徴とする請求項第1項に記載の情報処理装置。

【請求項4】 前記第2の制御手段は、音声入力手段から入力された音声情報の認識結果によって制御することを特徴とする請求項第1項に記載の情報処理装置。

【請求項5】 前記第2の制御手段が、応答文の出力再開を行うことを決定した場合、中断された位置より遡る方向に沿って区切り位置をサーチし、当該サーチした位置から出力を再開することを特徴とする請求項第1項に記載の情報処理装置。

【請求項6】 前記区切り位置は、一文節、一アクセント、一呼吸段落、或いは文頭が含まれることを特徴とする請求項第1項に記載の情報処理装置。

【請求項7】 音声入力手段と、音声認識手段と、前記音声入力手段に対して応答する出力手段を備えた情報処理装置の制御方法において、

前記音声入力手段から入力された音声を認識し、音声入力に対応する文を生成する文生成工程と、  
該文生成工程で生成された文に対応する出力文を構成し、前記出力手段を制御して外部に出力する出力制御工程と、

該出力制御工程で出力中、前記音声入力手段からの入力を監視し、音声入力があった場合に、当該出力を一時的に停止する第1の制御工程と、

入力された音声に応じて、出力途中であった応答文を再開するか、中断するかを制御する第2の制御工程とを備えることを特徴とする情報処理装置の制御方法。

【請求項8】 前記出力手段は、音声出力手段及び対話文を表示する表示手段の両方を含むことを特徴とする請求項第7項に記載の情報処理装置の制御方法。

【請求項9】 前記第2の制御工程は、音声入力手段か

2

ら入力された音声情報の長さによって制御することを特徴とする請求項第7項に記載の情報処理装置の制御方法。

【請求項10】 前記第2の制御工程は、音声入力手段から入力された音声情報の認識結果によって制御することを特徴とする請求項第7項に記載の情報処理装置の制御方法。

【請求項11】 前記第2の制御工程が、応答文の出力再開を行うことを決定した場合、中断された位置より遡る方向に沿って区切り位置をサーチし、当該サーチした位置から出力を再開することを特徴とする請求項第7項に記載の情報処理装置の制御方法。

【請求項12】 前記区切り位置は、一文節、一アクセント、一呼吸段落、或いは文頭が含まれることを特徴とする請求項第7項に記載の情報処理装置の制御方法。

## 【発明の詳細な説明】

【0001】

【産業上の利用分野】本発明は情報処理装置及びその制御方法、詳しくは発声に対して応答する情報処理装置及びその制御方法に関するものである。

【0002】

【従来の技術】最近、音声対話装置が用いられるようになってきた。図3を用いて、従来の音声対話方式について説明する。まず、マイク等に向かってユーザが図3(1)のように発声したとする。マイクに入力された音声は、A/D変換部2によってアナログ信号からデジタル信号に変換され、そのデジタル信号は音声認識部3によって図3(1)のような日本語文に変換される。この日本語文は対話管理部4によって解釈され、それに対応する図3(2)のような応答文が作成される。この応答文は音声合成部6によって単語に分割され、読みとアクセントが付加され、音韻パラメータと韻律パラメータとのデジタル信号に変換され、D/A変換部7によってアナログ信号に変換されて、スピーカ8によって音声として出力される。以上のような音声入力と音声出力との繰り返しによ、ユーザとシステムとの対話が行なわれていた。

【0003】

【発明が解決しようとする課題】上述した従来の音声対話方式では、音声出力中にユーザが音声を入力する際、ユーザの入力を無視して合成が続いたり、ユーザの入力により合成がストップしたりして、対話の自然性がなかった。

【0004】

【課題を解決するための手段】及び

【作用】本発明は、上記問題点に鑑みなされたものであり、音声入力による対話を自然な状態にさせると共に、尚且つ、ユーザにとっては必要な情報を迅速に得ることを可能ならしめる情報処理装置及びその制御方法を提供しようとするものである。

【0005】この課題を解決するため、例えば本発明の情報処理装置は以下の構成を備える。すなわち、音声入力手段と、音声認識手段と、前記音声入力手段に対して応答する出力手段を備えた情報処理装置において、前記音声入力手段から入力された音声を認識し、音声入力に対応する文を生成する文生成手段と、該文生成手段で生成された文に対応する出力文を構成し、前記出力手段を制御して外部に出力する出力制御手段と、該出力制御手段で出力中、前記音声入力手段からの入力を監視し、音声入力があった場合に、当該出力を一時的に停止する第1の制御手段と、入力された音声に応じて、出力途中であった応答文を再開するか、中断するかを制御する第2の制御手段とを備える。

【0006】また、本発明の実施態様に従えば、前記出力手段は、音声出力手段及び対話文を表示する表示手段の両方を含むことが望ましい。これによって、入力文と応答文の両方が表示されるので、音声出力された内容を確実に確認することが可能になる。

【0007】また、第2の制御手段は、音声入力手段から入力された音声情報の長さによって制御しても良い。この結果、音声の長さのみで良いので、制御を簡単にすることが可能になる。

【0008】また、第2の制御手段は、音声入力手段から入力された音声情報の認識結果によって制御するようにしても良い。この結果、ユーザの意図を察知して、対応する処理を確実に行うことが可能になる。

【0009】また、第2の制御手段が、応答文の出力再開を行うことを決定した場合、中断された位置より遡る方向に沿って区切り位置をサーチし、当該サーチした位置から出力を再開することが望ましい。特に、その区切り位置は、一文節、一アクセント、一呼吸段落、或いは文頭が含まれるようにすることで、出力が途中であっても、その出力文を正しく、且つ、自然な状態で出力することが可能になる。

【0010】

【実施例】以下、添付図面に従って本発明に係る実施例を詳細に説明する。

【0011】図2は、本実施例の音声対話装置の構成を示すブロック図である。この装置は、マイク1、A/D変換部、音声認識部3、対話管理部4、会話内容等を表示する表示部5、音声合成部6、D/A変換部7、そしてスピーカ8から構成されている。また、CPU9は、その内部の主メモリに格納されたプログラムに従って上記各構成要素を制御する。

【0012】CPU9の処理手順としては、図1の通りであり、音声合成出力開始ステップS11と、音声入力が存在するかどうかを判定するステップS12と、音声合成出力を継続するステップS13と、音声入力終了したかどうかを判定するステップS14と、音声合成出力を一時中断するステップS15と、音声入力存続期間

がある閾値より大きい小さいかを判定するステップS16と、音声合成出力を終了するステップS17となる。

【0013】次に、このように構成された本実施例の音声対話方式の動作について、図1および図4を参照しながら説明する。

【0014】まず、ユーザが図4の手順(1)のように発生したとする。すると、ステップS11によって図4の手順(2)のように音声合成出力が開始される。音声合成出力中にも音声認識部3は入力を受け付け、その結果は随時対話管理部4に送られており、対話管理部4はステップS12によって音声入力が存在するかどうかを判定し、入力が存在しなければステップS13によって音声合成出力を継続し、ステップS12に戻る。ステップ14によって音声入力終了したと判定された場合は、ステップ16に移る。ステップ16によって音声入力存続期間がある閾値Thより小さいと判定された場合には、ステップ13によって音声合成出力を継続し、ステップ12に戻る。ステップ16によって音声入力存続期間がある閾値Thより大きいと判定された場合には、ステップ17に移って音声合成出力を終了する。

【0015】いま、図4の手順(2)のようにシステムが「神戸市立須磨海浜水族館の説明は、24000平方mの敷」まで音声合成出力した段階で、図4の手順(3)のようにユーザが「えっ」と発生したとする。すると、ステップS12によって音声入力が存在すると判定され、ステップS14とステップS15とによって音声合成出力が一時中断される。

【0016】ここでは、ステップ16によって「えっ」の存続期間がThより小さいと判定されたとする。すると、ステップS17によって、図4の手順(4)のように「地に水族園本館、ラッコ館など…」と音声合成出力が継続される。さらに、図4手順(4)のようにシステムが「…の大群が泳ぎ、ラッ」まで音声合成出力した時点で、図4の手順(5)のようにユーザが「ありがとう」と発生したとする。すると、ステップ12によって音声入力が存在すると判定され、ステップS14とステップS15とによって音声合成出力が一時中断される。ここではステップ16によって「ありがとう」の存続期間がThより大きいと判定されたとする。すると、ステップ17に移って音声合成出力を終了する。以上の対話状況は図4のように、表示部5の画面上に表示される。

【0017】以上説明したように本実施例によれば、応答文に基づく音声合成出力中に、ユーザからの発声を検出したとき、その音声合成出力を一時停止し、その発声内容(上記実施例では存続期間)に従って、音声合成出力を継続するか、または、中断するかを決定する。従って、上記の如く、ユーザがそれ以上の情報の必要性がないと判断した場合には、「ありがとう」等発声することで、それ以降の音声合成を止めさせることが可能にな

5

り、場合によっては次の問い合わせを即座に行うことも可能になる。

【0018】[第2の実施例の説明]次に、本発明の第2の実施例について述べる。

【0019】図5を参照すると、本実施例の音声対話方式における処理は、音声合成出力開始ステップS51、音声入力が存在するかどうかを判定するステップS52、音声合成出力を継続するステップS53、音声入力

が終了したかどうかを判定するステップS54、音声合成出力を一時中断するステップS55、音声入力存続期間がある閾値より大きい小さいかを判定するステップS56、音声合成出力の継続位置を決定するステップS57、そして、音声合成出力を終了するステップS58とからなる。なお、本第2の実施例の上記各ステップはCPU9内の主メモリに予め記憶されているものである。

【0020】次に、このように構成された本実施例の音声対話方式の動作について、図5および図6を参照しながら説明する。

【0021】まず、ユーザが図6の手順(1)のように発声したとする。これに対してシステムが図6の手順(2)のように「神戸市立須磨海浜水族館の説明は、24000平方mの敷」まで音声合成出力した段階で、図6の手順(3)のようにユーザが「えっ」と発声し、その存在期間がThより小さいと判定されたとする。すると、ステップS57によって音声合成出力の再開位置が中断された文節の一文節前の文節の先頭位置に設定される(但し、再開位置はアクセント句前または一呼吸段落前など、区切りの良い位置ならばどこでもよい)。

【0022】次に、ステップS53に移って図6の手順(4)のようにシステムが「24000平方mの敷地に水族園本館…」と音声合成出力を継続する。さらに、図6の手順(4)のようにシステムが「…の大群が泳ぎ」まで音声合成出力した時点で、図6(5)のようにユーザが「ありがとう」と発生したとする。するとステップS52によって音声入力が存在すると判定され、ステップS54とステップS55によって音声合成出力が一時中断される。ここでは、ステップS56によって「ありがとう」の存続時間がThより大きいと判断されたとする。すると、ステップS58に移って音声合成出力を終了する。以上の対話状況は図6のように表示部5の画面上に表示される。

【0023】以上の結果、本第2の実施例によれば、音声合成出力中に、一時的にそれを中断した場合には、その中断位置からではなく、その直前の文節や、アクセント等聞き取り易い位置から再開するので、上記第1の実施例と比較し、ユーザフレンドリィになる。

【0024】[第3の実施例の説明]次に、本発明の第3の実施例について述べる。

【0025】図7を参照すると、本実施例の音声対話方

6

式における処理は、音声合成出力開始ステップS71、音声入力が存在するかどうかを判定するステップS72、音声合成出力を継続するステップS73、音声入力

が終了したかどうかを判定するステップS74、音声合成出力を一時中断するステップS75、音声認識結果が肯定的か否定的かを判定するステップS76、音声合成出力の継続位置を決定するステップS77、そして、音声合成出力を終了するステップS78とからなる。

【0026】次に、このように構成された本実施例の音声対話方式の動作について、図7および図8を参照しながら説明する。

【0027】まず、ユーザが図8の手順(1)のように発生したとする。これに対してシステムが図8の手順(2)のように「神戸市立…など7館が点在」と応答したとする。この時点でユーザが図8の手順(3)のように「はい」と応じると、ステップ72によって音声入力が存在すると判定され、ステップ75によって音声合成出力が一時中断し、ステップ76によって音声入力「はい」の認識結果が肯定的か否定的かを判定する。ここでは「はい」「いいえ」「わかりました」等は肯定的であるとみなして、ステップ77に移って音声合成出力の継続位置が決定される。ここでは、文の終了時点なので区切りが良いとみなし、ステップ73に移って音声合成出力が継続され、図8の手順(4)のように「間口25m…の大群が泳ぎ、ラッコの」と音声合成出力が続けられる。この時点でユーザが図8の手順(5)のように「終了してください」と肯定的でない発生をするとステップ76によって認識結果が否定的であると判定されて、ステップ78に移り音声合成出力を終了する。以上の対話状況は図8のように、表示部5の画面上に表示される。

【0028】以上説明したように本実施例によれば、随時入力を受け付ける音声認識手段と、入力に対する応答を音声で出力する音声合成手段と、入力/出力を管理する対話管理手段と、生成する応答文の音声合成による応答をユーザの入力に応じてコントロールする手段を設けることにより、ユーザとシステムとの対話の自然性が改善される。

【0029】尚、本発明は、複数の機器から構成されるシステムに適用しても、1つの機器から成る装置に適用しても良い。また、本発明はシステム或は装置にプログラムを供給することによって達成される場合にも適用できることは言うまでもない。

【0030】

【発明の効果】以上説明したように本発明によれば、音声入力による対話を自然な状態にさせると共に、尚且つ、ユーザにとっては必要な情報を迅速に得ることが可能になる。

【0031】

【図面の簡単な説明】

【図1】本発明の第1の実施例の音声対話方式の処理を



示す流れ図である。

【図2】本実施例の音声対話方式が適用される音声対話装置の構成の一例を示すブロック図である。

【図3】従来の音声対話方式の動作例を示す図である。

【図4】本発明の第1の実施例の動作例を示す図である。

【図5】本発明の第2の実施例の処理を示す流れ図である。

【図6】本発明の第2の実施例の動作例を示す図である。

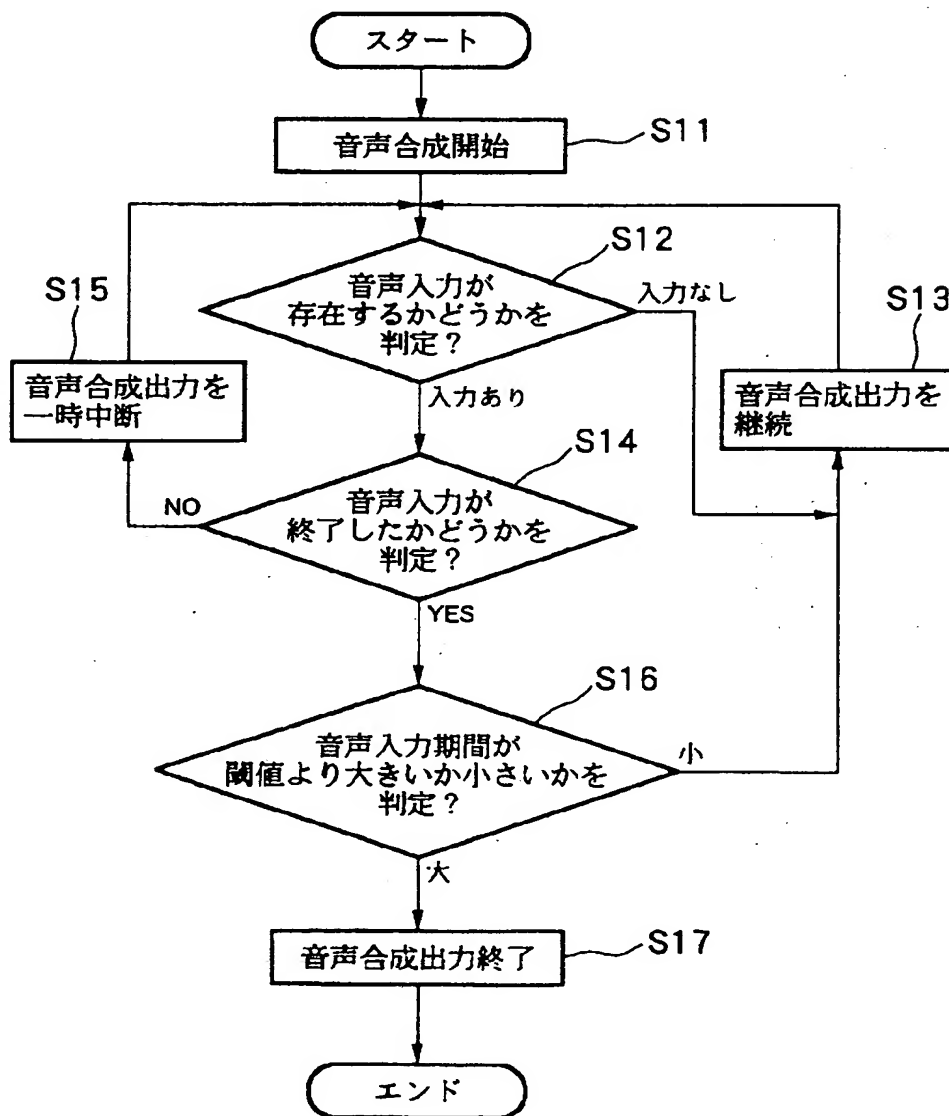
【図7】本発明の第3の実施例の処理を示す流れ図である。

【図8】本発明の第3の実施例の動作例を示す図である。

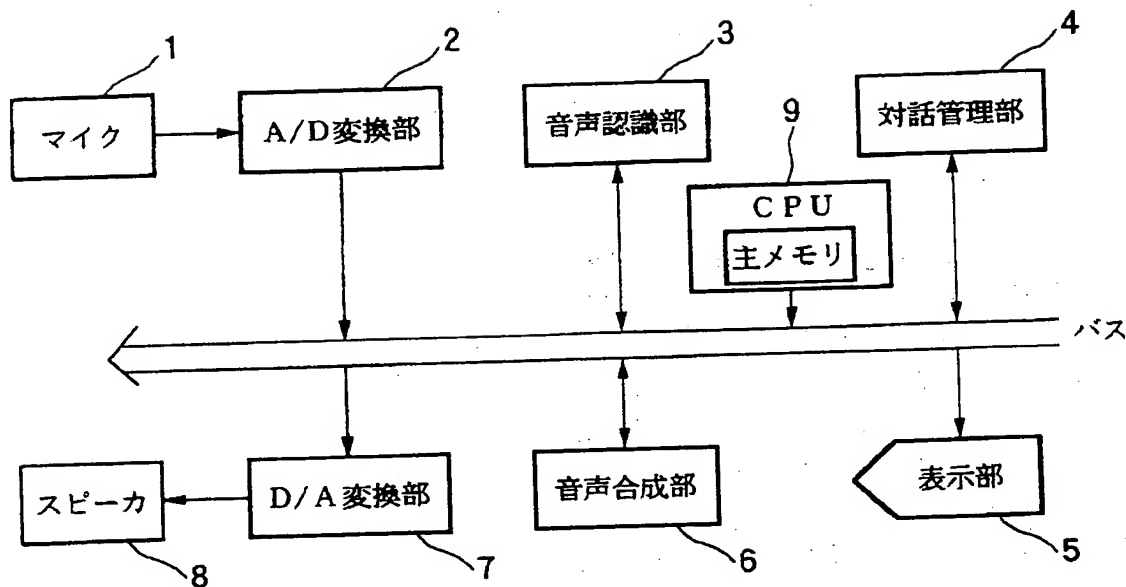
【符号の説明】

- 1 マイク
- 2 A/D変換部
- 3 音声認識部
- 4 対話管理部
- 5 表示部
- 6 音声合成部
- 10 7 D/A変換部
- 8 スピーカ

【図1】



【図2】



【図3】

- (1) USER : 神戸市立須磨海浜水族園の説明は。  
 (2) SYSTEM : 神戸市立須磨海浜水族園の説明は、24000平方mの敷地に水族園本館、ラッコ館など7館が点在。間口25m、高さ4mの波の大水槽には、大きな外洋性サメや魚の大群が泳ぎ、ラッコの岩盤ではシアトルから来たかわいいラッコたちが愛敬をふりまいている。

【図4】

- (1) USER : 神戸市立須磨海浜水族園の説明は。  
 (2) SYSTEM : 神戸市立須磨海浜水族園の説明は、24000平方mの敷地に水族園本館、ラッコ館など7館が点在。間口25m、高さ4mの波の大水槽には、大きな外洋性サメや魚の大群が泳ぎ、ラッコの岩盤ではシアトルから来たかわいいラッコたちが愛敬をふりまいている。  
 (3) USER : えっ  
 (4) SYSTEM : 地に水族園本館、ラッコ館など7館が点在。間口25m、高さ4mの波の大水槽には、大きな外洋性サメや魚の大群が泳ぎ、ラッコの岩盤ではシアトルから来たかわいいラッコたちが愛敬をふりまいている。  
 (5) USER : ありがとう

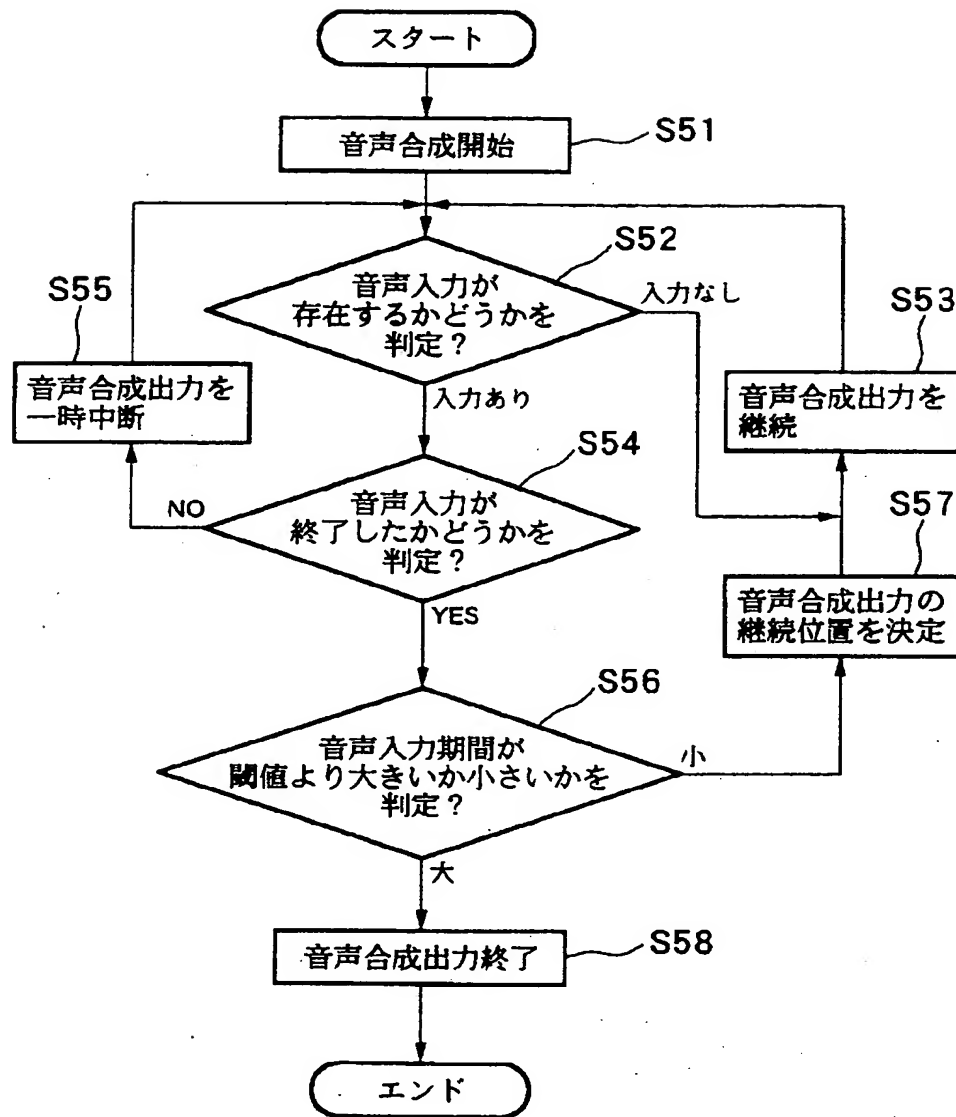
【図6】

- (1) USER : 神戸市立須磨海浜水族園の説明は。  
 (2) SYSTEM : 神戸市立須磨海浜水族園の説明は、24000平方mの敷地に水族園本館、ラッコ館など7館が点在。間口25m、高さ4mの波の大水槽には、大きな外洋性サメや魚の大群が泳ぎ、ラッコの岩盤ではシアトルから来たかわいいラッコたちが愛敬をふりまいている。  
 (3) USER : えっ  
 (4) SYSTEM : 24000平方mの敷地に水族園本館、ラッコ館など7館が点在。間口25m、高さ4mの波の大水槽には、大きな外洋性サメや魚の大群が泳ぎ、ラッコの岩盤ではシアトルから来たかわいいラッコたちが愛敬をふりまいている。  
 (5) USER : ありがとう

【図8】

- (1) USER : 神戸市立須磨海浜水族園の説明は。  
 (2) SYSTEM : 神戸市立須磨海浜水族園の説明は、24000平方mの敷地に水族園本館、ラッコ館など7館が点在。  
 (3) USER : はい  
 (4) SYSTEM : 間口25m、高さ4mの波の大水槽には、大きな外洋性サメや魚の大群が泳ぎ、ラッコの岩盤ではシアトルから来たかわいいラッコたちが愛敬をふりまいている。  
 (5) USER : 終了してください

【図5】



【図7】

